

(43) Date of A Publication 18.06.1997

(21) Application No 9525791.1

(22) Date of Filing 16.12.1995

(71) Applicant(s)
International Business Machines Corporation
(Incorporated in USA - New York)
Armonk, New York 10504, United States of America

(72) Inventor(s)
Ronald John Bowater
Mervyn Anthony Steton
Gerard Richter

(74) Agent and/or Address for Service
S R Davies
IBM UK Ltd, Intellectual Property Dept, Hursley Park,
Winchester, Hampshire, SO21 2JN, United Kingdom

(51) INT CL⁸
H04B 3/20

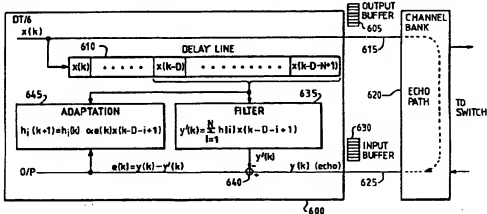
(52) UK CL (Edition O)
H4R RLES
U1S S2215

(56) Documents Cited
GB 2109209 A GB 2109207 A

(58) Field of Search
UK CL (Edition O) H4R RLES RLS
INT CL⁸ H04B
Online: WPI, JAPO, INSPEC

(54) System and method for echo cancellation

(57) The invention relates to a system and method for echo cancellation within a voice processing system e.g. voice mail. Echo signals received by a voice processing system can be misinterpreted as commands from a user thereof thereby causing the system to operate incorrectly. Establishing echo cancellation coefficients, especially on a call by call basis, demands a significant amount processing resource. Conventionally, the echo cancellation coefficients have been calculated using a convolution operation which is again very processor intensive. The present invention calculates and uses echo cancellation coefficients using a simple subtraction technique and stores the results for repeated later use in cancellation thereby obviating the need to establish echo cancellation parameters on a call-by-call basis and reducing the computation overhead in performing echo cancellation. A signal $n(k)$ is applied via a line 615 and a channel bank including hybrids producing local echoes to a telephone switching network. An estimate of the echo signal $y^*(k)$ is computed from echo cancellation coefficients $h(k)$ derived using delay line 610 and filter 635. The estimated echo signal $y^*(k)$ is subtracted at 640 from actual echo signal $y(k)$ and used to adaptively modify the echo cancellation coefficients.

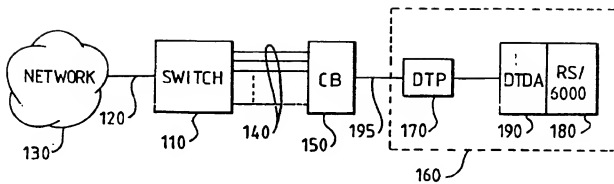
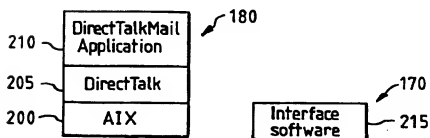


D : FIXED DELAY (20ms PACKETIZATION DELAY, FLIP_FLOP)
N : FILTER LENGTH 125 μ s PER TAP (FOR N=32 : 4ms)
or : ADAPTATION CONTROL, DEPEND ON $-\frac{N}{E(x^2)}$ SIGNAL POWER TO BE FIXED DURING TEST

FIG. 6

GB 2 308 283 A

BEST AVAILABLE COPY

**FIG. 1****FIG. 2**

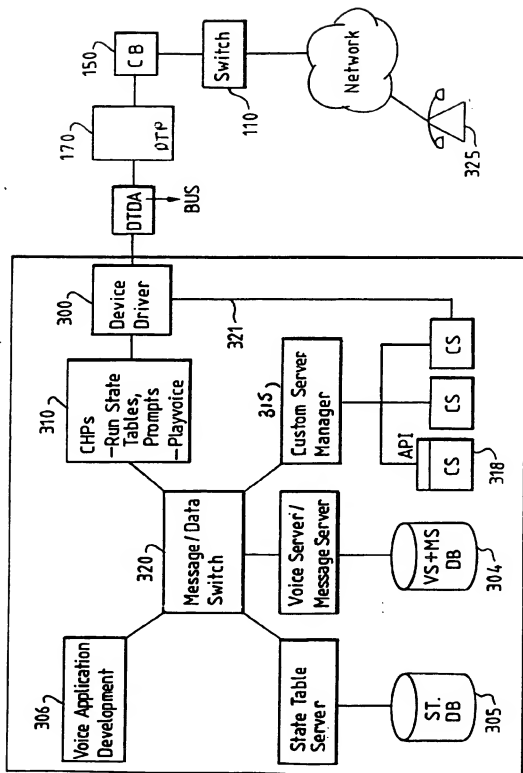
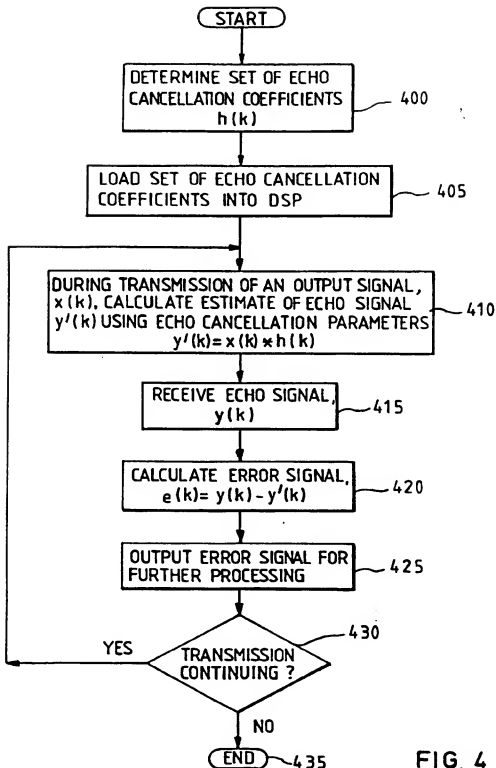


FIG. 3

FIG. 4

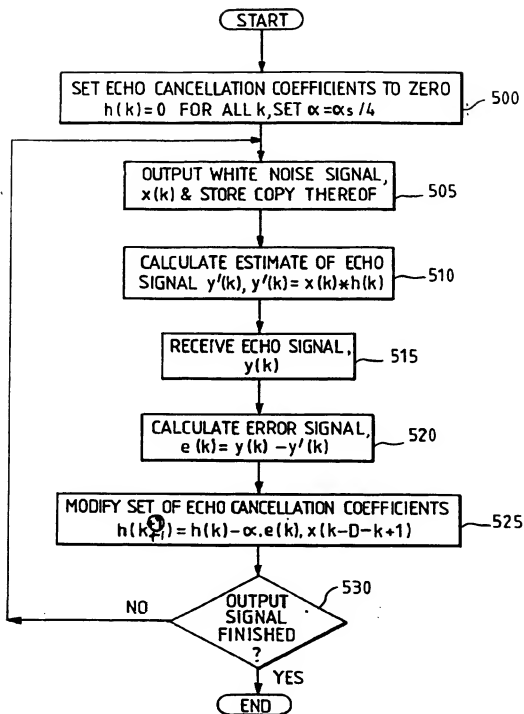
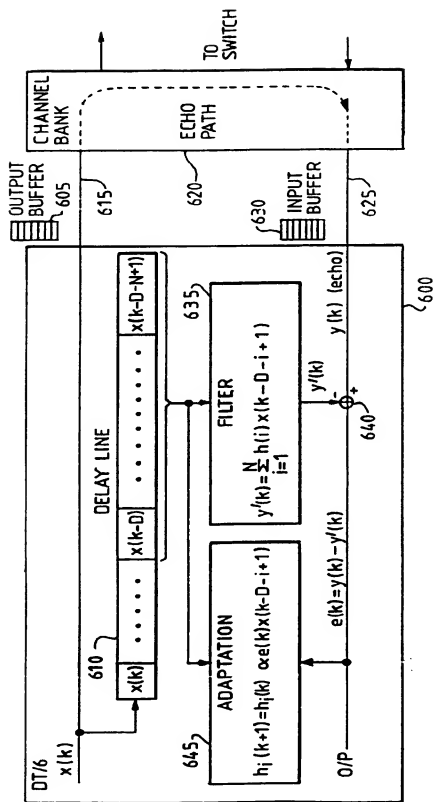


FIG. 5



D : FIXED DELAY (20ms PACKETIZATION DELAY, FLIP-FLOP)

N : FILTER LENGTH 125 μ s PER TAP (FOR N=32 : 4ms)

α : ADAPTATION CONTROL, DEPEND ON $-\sum_{i=1}^N e^2(x^2)$ SIGNAL POWER TO BE FIXED DURING TEST

FIG. 6

SYSTEM AND METHOD FOR ECHO CANCELLATION

The present invention relates to a system and method for echo cancellation, more particularly, to a such a system and method for use with a voice processing system.

Voice processing systems, which are well-known in the art (see for example "Voice Processing", by Walt Teschner, published by Artech House), perform a variety of functions, the most common of which is voice mail (also known as voice messaging), whereby callers who cannot reach their intended addressee can instead record a message for them for subsequent retrieval.

An example of a voice response system is the IBM CallPath DirectTalk/6000 product as described in "IBM CallPath DirectTalk/6000 General Information and Planning" and "IBM CallPath DirectTalk/6000 Voice Application Development" (IBM, DirectTalk, DirectTalk/6000 and CallPath are trade marks of International Business Machines Corporation).

Voice response systems enable users thereof to access information using a conventional telephone. The interaction between the users and the system comprises various voice prompts output by the system and responses thereto as inputs, such as DTMF tones or voice, by the user. Voice response systems are used by service providers, such as banks, to fully or partially automate telephone call answering or responding to queries. Typically a voice response system provides the capability to play voice prompts comprising recorded voice segments or speech and to receive responses thereto. The prompts may be organised in the form of voice menus invoked by state tables. A state table contains commands which can access and play a voice segment or synthesise speech from given text. The prompts are usually part of a voice application which is designed to, for example, allow a customer to query information associated with their various banks accounts. While the voice prompt is being played by the voice processing system, a digital signal processor (DSP) monitors an input line for input data, such as a DTMF tone or audio data. EP-A-622964 discloses a system and method suitable for detecting voice activity or DTMF tones on an incoming telephone line. If, for example, a DTMF tone is detected during the output of the current voice prompt it is terminated by informing the process responsible for obtaining the digitised audio data units from the voice/message database of the interruption.

Conventionally, a digital voice processing system is connected to a remote telephone via a communication network, which typically involves other switches, hybrids and a plurality of transmission media. The hybrids within the transmission media used to establish the call are a source of echo signal. If a digital voice processing system, such as DirectTalk/6000, is being used it is often necessary to connect the voice processing unit to the switch via a channel bank because most of the functionality, such as call transfer or call forwarding, of switches is generally made available for use by the voice processing via analog lines. The channel bank may comprise a plurality of hybrids for converting four-wire lines from the voice processing unit to a two-wire lines of the switch. The channel bank is therefore also a source of echo signals. The echo signals, if of sufficient magnitude to be acted upon by the DSP monitoring incoming signals, can interfere with the correct operation of the voice processing system.

As mentioned above a DSP monitors the input lines to detect incoming signals such as audio data or DTMF tones. If the echoes are of sufficient magnitude they can be mistaken for incoming audio data or DTMF tones and thereby interfere with the correct operation of the voice processing system. Any such interference may, for example, cause a currently playing voice prompt to be prematurely terminated, or reduce the accuracy of voice recognition, for example, the echo may be erroneously interpreted as a voice signal from a caller.

Hence, voice processing systems conventionally include complex methods of eliminating echo signals present on an incoming signal. US 5,164,989 discloses an echo signal cancellation method and apparatus. Each time a telephone call is established, the echo characteristics of the connection are determined using a training sequence. The echo signal of the training sequence is recorded and stored. The stored sequence is convolved with the time-domain inverse of the training sequence to produce a function which approximates the transfer function or impulse response of the transmission line. This function is then used for subsequent echo cancellation.

The determination of echo signal cancellation parameters according to USA 5,164,989 is a very processor intensive operation and represents a significant drain on the DSP resources of a voice processing system as it involves computing the convolution of two digital signals. The DSPs which would have ordinarily been used for DTMF detection or voice

recognition are instead utilised in characterising the call connection or determining the characteristics of the transmission medium. Furthermore, within voice processing system, a single DSP may be responsible for processing the input data from a plurality of input lines and may be, for example, attempting to supervise or perform voice recognition for one input line while concurrently determining echo signal cancellation parameters for another line over which a call has recently been established. Still further, the determination of echo cancellation coefficients during a call has the disadvantage that any training sequence may be audible by the caller.

Accordingly, the present invention provides a method for performing echo cancellation, within a voice processing system during transmission of an audio signal over a communication network, said communication network comprising a local source of echoes, said method comprising the steps of initially determining a set of echo cancellation coefficients suitable for cancelling locally generated echoes, and performing echo cancellation using said set of echo cancellation coefficients for all subsequent transmissions over said communication network.

The same coefficients can be used for echo cancellation for all calls without the need to characterise the transmission media on a per call basis because the present invention is primarily concerned with reducing the effect of locally generated echo signals produced, for example, by the hybrids employed within the local channel bank to which the voice processing system may be connected. The local characteristics of the transmission medium between the voice processing system and the channel bank are fixed and therefore need to be determined only once. Hence the same echo cancellation coefficients can be used for all calls thereby obviating the need to the perform computationally intensive initial step of characterisation of the transmission medium for all calls. The amount of overall processing by a DSP to perform echo signal cancellation is thereby reduced and hence the resources of the voice processing system can be more effectively utilised to perform other functions such as DTMF tone detection or voice recognition.

The present invention does not completely eradicate echo signals received by the voice processing system, but reduces the overall echo signals received to such a level that a DSP monitoring the input line is less likely to erroneously interpret an echo signal as a data for which processing is required, that is, the echo signal is reduced to below a

predeterminable level. It is the locally generated echo signals which represent the most significant contribution to the overall received echo signals. Generally, the more remote a source of an echo is, relative to the voice processing system, the less significant the contribution thereof to the overall received echo signal. In effect, remotely generated echoes are assumed to be relatively insignificant and ignored or cancelled by echo cancellers within the communication network. Echo cancellation coefficients are determined which predominantly reduce the locally generated echo signals. A source of locally generated echoes is typically the first hybrid used to connect the voice processing system to a communication network or to a switch. The delay between the transmission of a signal and receiving a locally generated echo thereof is typically less than eight milliseconds.

Conventionally, the echo cancellation coefficients have been calculated on the basis of a characterisation of the whole of the transmission link from the voice processing system to the telephone. Most echo cancellers are designed to cancel echo of signals which are received up to thirty-two milliseconds after the transmission of the outgoing signal from which the former is derived. Characterisation and subsequent processing of such echo signals requires a very long digital filter length and invariably involves a significant amount of processing power.

Suitably, an embodiment of the present invention provides a method, wherein the digital filter length is arranged to store less than eight milliseconds worth of digital data.

Such a short filter length can be used because the present invention is only concerned with reducing the impact of the relatively more significant locally generated echo signals.

As mentioned above, using discrete time domain convolution to determine the impulse response or transfer function of a transmission medium to which a voice processing system is connected is a very computationally intensive task. Performance of any such convolution again represents a significant drain upon the limited processing resources of a voice processing system.

Accordingly, an embodiment of the present invention provides a method wherein said step of initially determining a set of echo cancellation coefficients comprises the steps of

- (a) setting all echo cancellation coefficients to a value of zero,
- (b) transmitting an output signal $(x(k))$ to said transmission medium (515),
- (c) storing a copy of said output signal $(x(k))$ in a buffer,
- (d) receiving from said transmission medium (525) an echo signal $(y(k))$ representative of a locally distorted version of said output signal $(x(k))$,
- (e) generating an estimate $(y'(k))$ of said echo signal using said copy of said output signal, and
- (f) modifying said echo cancellation coefficients according to the difference between said echo signal $(x(k))$ and said estimate $(y'(k))$ thereof.

Modifying the echo cancellation coefficients according to the difference between said actual echo signal and said estimate of thereof involves very simple arithmetic and as such imposes less of a demand upon the processing resources of the voice processing system.

In a preferred embodiment, the steps (b) to (f) are repeated a number of times until said modified echo cancellation coefficients converge to stable values. With each iteration of steps (b) to (f) the echo cancellation coefficients gradually converge to said stable values. Each iteration effectly produces a more progressively refined, complete set of echo cancellation coefficients, $h_i(k)$, where i takes values between 1 and N , the latter being the filter length, and k is the k th iteration or a value representative of a point in the discrete-time domain.

It will be appreciated that the order of execution steps (d) and (e), in a particular is immaterial. In other embodiments, steps (d) and (e) may be executed substantially concurrently. In a still further embodiment the calculation of the estimate of the echo signal may be

commenced before processing of the incoming signal as there invariably exists a small delay between transmission of an outgoing signal and the receipt of echoes derived therefrom. The delay can be effectively utilised to calculate, or at least partially calculate, said estimate of said echo signal.

Furthermore, by calculating an estimate of the echo signal while awaiting the receipt of the actual echo signal, the available processing time of the voice processing system is more efficiently utilised. The DSPs of the voice processing system do not have to await the receipt of an incoming signal before calculations relating to echo cancellation can be commenced.

The present invention also provides a system for performing echo cancellation, within a voice processing system during transmission of an audio signal over a communication network between said audio signal and a telephone, said communication network comprising a local source of echoes, said system comprising means for initially determining a set of echo cancellation coefficients suitable for cancelling locally generated echoes, and means for performing echo cancellation using said set of echo cancellation coefficients for all subsequent transmissions over said communication network.

Embodiments of the invention will now be described in detail, by way of example only, with reference to the following drawings:

figure 1 is a simple block diagram showing a voice processing system connected to a telephone switch via a channel bank,

figure 2 is a simple block diagram of the main software components of a DirectTalk/6000 system,

figure 3 shows schematically echo cancellation according to an embodiment,

figure 4 illustrates schematically initialisation of the echo cancellation coefficients,

figure 5 shows a schematic flow diagram of an embodiment.

Figure 1 is a simple block diagram showing a switch 110 which exchanges telephony signals with the external telephone network 130 over digital trunk line 120. The switch is logically divided into two halves, namely, a tie-line side, which is of no relevance to the present invention and will therefore not be described further, and a station side. The station side provides a plurality of two-wire analog lines via which the functionality of the switch is made available. A voice processing system desiring to take advantage of the available functionality must be connected to the analog lines 140 via a channel bank 150. The channel bank 150 conventionally contains a plurality of hybrids which allow the connection of the two-wire lines of the switch to the four-wire lines of the voice processing system. The hybrids within the channel bank 150 are a source of locally generated echo signals which may be received by and adversely impact the operation of the voice processing system. In a current implementation, the voice processing system is a DirectTalk/6000 system (ie runs the DirectTalk/6000 software), but the same principles apply whatever voice processing system is being used.

The DirectTalk/6000 system comprises two main hardware components, a digital trunk processor 170, and computer workstation 180, which in the case of the DirectTalk/6000 system is a RISC System/6000. Also shown is an adapter card 190 (DTDA), which provides an interface between the RISC System/6000 and the telephone interface module. Note that in many voice processing systems, the telephone interface module is incorporated into the adapter card for direct attachment to the computer workstation. The DirectTalk/6000 system (software plus hardware) is available from IBM Corporation, and is described more fully in IBM Callpath DirectTalk/6000 General Information and Planning (reference number GC22-0100-03) and other manuals mentioned therein, also available from IBM. As stated above, although the invention is being described with reference to the DirectTalk system, it can be utilised in many other environments for which echo cancellation is required, such as within modems or voice recognition applications or the like.

Figure 2 is a simple block diagram of the main software components of a DirectTalk/6000 system. Running on the RISC System/6000 is first of all the operating system 200 for the workstation, which in the present case is AIX, and then the DirectTalk/6000 software 205 itself. Finally, also running on the RISC System/6000 workstation is an application 210, in this case DirectTalkMail, which interacts with the operating system

and the DirectTalk/6000 software to provide the desired voice mail function. Various routines 215 also run within the digital trunk processor 170. These routines are downloaded from the RISC System/6000 onto the telephone interface module when the telephone interface module is enabled, and handle items such as detection of tones, silence, voice, generation of tones.

Figure 3 is a schematic diagram of the main components of a DirectTalk/6000 system. Only those components relevant to an understanding of the present invention will be described; further details can be found in the above-mentioned manuals. The first set of components run on the RISC System/6000 workstation 180 and comprise a device driver 300 which is used to interact via the adapter card 190 (Dual Trunk Digital Adapter, DTD A) with the digital trunk processor 170. A state table 305 provides the program control of applications executing in the DirectTalk/6000 system (ie in developing an application, the custom creates a set of state tables). The channel processor (CHP) 310 contains the code which performs the actions specified by the state tables 305. A custom server manager 315 allows external connections into and out of the DirectTalk/6000 system. The custom server 318 can operate in one of two modes. Firstly, it can perform simple functions as requested by a state table and return data as appropriate. Secondly, it can fetch voice data from the voice segment database 304 via the message/data switch 320, process that data and then feed it directly to the device driver 300 via the custom server voice services interface communication 321. The above is described in more detail DirectTalk/6000 Voice Application Development Guide SC22-0102-03, specifically under the routine CA_Play_Voice_Stream.

DTMF tones are detected by one of the DSPs in the DTP 170 implementing an appropriate digital filter. The DTP 170 informs the device driver 300 that a DTMF tone has been detected and the DTMF key to which the tone corresponds. The device driver then interrupts the output of the audio data by informing the custom server responsible for obtaining the digitised audio data units from the voice/message database.

Upon installation of a voice processing system, an application is run on the CHP 310 to determine the local echo characteristics which result from the connection of the voice processing system to a local channel bank 150. According to the embodiment realised using DirectTalk/6000, a call connection is established to a remote telephone. As it is the local echo characteristics which are to be determined, the

call connection can be to any remote telephone. Alternatively, an embodiment can be realised in which the characterisation is performed passively, that is without their being established an actual call connection. The application determines the echo cancellation coefficients, $h(k)$, as follows. An output signal, $x(k)$, is transmitted by the voice response system to the transmission medium to be characterised, either passively or by way of a call connection. The output signal can be generated in advance by a further application executing on the CHP 310 and stored in one of the data bases 304, 350. Alternatively, the output can be generated in real-time by a custom server 318 at the instigation of an application. An echo signal, $y(k)$, is received from the transmission medium. The echo signal represents a delayed and distorted version of $x(k)$. The extent of the distortion is dependent upon the transfer function representative of the path taken by the output signal, $x(k)$, between being output and subsequently received by the voice processing system. Echo cancellation involves computing an estimate of the echo signal, $y'(k)$, and subtracting that estimate from the actual echo signal, $y(k)$. The estimate of the echo signal, $y'(k)$, is representative of the following

$$y'(k) = h(1)x(k-D) + h(2)x(k-D-1) + \dots + h(N)x(k-D+N-1),$$

that is the convolution of the output signal and the transfer function of the transmission medium to be characterised. As mentioned above, $h(i)$ are the set of echo cancellation coefficients which represent the transfer function of the transmission path.

During initialisation, the coefficients are derived in an incremental manner as follows. An error signal, $e(k)$, is calculated which is representative of the difference between the estimate of the echo signal, $y'(k)$ and the actual echo signal, $y(k)$. Hence

$$e(k) = y(k) - y'(k).$$

Each coefficient, $h(i)$, is calculated using

$$h(i+1) = h(i) - \alpha \cdot e(k) \cdot x(k-D-i+1).$$

where α is an adaption control, and D is the packetisation delay.

5 The echo cancellation coefficients converge to stable values after
a period of about 500 milliseconds. The rate of convergence is dependent
upon the value of alpha. The value of alpha, α , is determined using an
initial heuristic/empirical estimate, α_0 , which is close the
10 stability limit of convergence of the equation above. The value of α
depends upon factors such as the power of the transmitted signal and the
length of the filter. Once initially determined, the value of alpha
remains constant thereafter. It has been found that once a suitable
initial estimate of alpha has been determined, the value used for
15 convergence may be set to $\alpha = \alpha_0/4$. Convergence results in approximately
500 milliseconds using such a value of α . An improvement in the
convergence can be realised if a signal processing buffer sufficient to
accommodate a 1 second signal is utilised during initialisation.
Further, as the output or training signal, $x(k)$, has a constant power,
15 there is no need to adjust the value of α during the computation.

20 Although the above embodiment incrementally refines the echo
cancellation coefficients, a crude estimate thereof can be obtained by
executing a single pass through the above steps.

25 The packetisation delay, D , represents the delay incurred as a
consequence of placing the data representing the output signal into
packets for subsequent transmission over the transmission medium.
Although the packetisation delay, D , can be expected to vary from system
to system, it has been found to be approximately twenty milliseconds for
the DirectTalk/6000 system. Hence, the packetisation delay should be
accounted for during the calculation of the echo cancellation
coefficients and any subsequent echo cancellation.

30 A suitable output signal is white noise signal which can be
obtained by generating a random binary sequence. The white noise can be
generated in the form of a pseudo-random binary sequence.

35 The coefficients are then stored for subsequent loading into the
DSPs which perform input line monitoring for use in echo cancellation.
Accordingly, there is no need for the DSPs to characterise the
transmission line each time a call is established since the local line
echo characteristics remain substantially unchanged on a per call basis.

40 Referring to figure 4, there is shown a flow diagram illustrating
the calculation of the echo cancellation coefficients, and hence the

impulse response or transfer function of the transmission medium to be characterised. At step 400, the echo cancellation coefficients, $h[i]$, are determined. The range of values of i varies according to the desired length of the filter which in turn determines the which echoes are cancelled or the extent of the characterisation of the transmission medium. If a short filter length is used, those echoes which originate relatively locally will be cancelled. If a long filter length is used, the echoes cancelled will also include those which are generated at relatively remote distances from the voice processing system as the longer filter length characterises more of the transmission medium.

The cancellation of the relatively more significant locally generated echoes, having a delay of approximately eight milliseconds, requires a filter length of 64 samples, assuming the echo signalled is sampled at 8 kHz. Generally, the filter length is governed by the elapsed time between transmitting an output signal and receiving an echo thereof divided by the sampling period.

Having determined the echo cancellation coefficients, they are utilised as follows. At step 405 the echo cancellation coefficients are loaded into the DSPs responsible for performing echo cancellation during transmission of a signal, $x(k)$, and monitoring the echo signal for DTMF tones or other inputs such as voice from the caller. The DSP calculates an estimate, $y'(k)$, of the echo signal using the echo cancellation coefficients, $h(k)$, at step 410. An incoming signal, $y(k)$, is received by the voice processing system at step 415. The estimate of the echo signal, $y'(k)$ is subtracted from the incoming signal, $y(k)$, to form an error signal, $e(k)$ at step 420. At step 425, the error signal, $e(k)$, is used as the basis for further processing, such as determining whether or not the echo signal comprises signals, such as DTMF tones, other than echo signals. A test as to whether or not transmission of the output signal, $x(k)$, is continuing, and hence whether or not cancellation can be terminated, is made at step 430. If cancellation is still required echo cancellation continues from step 410. If cancellation is no longer required, then cancellation process terminates at step 435.

Referring to figure 5, there is shown in greater detail the step of determining the initial echo cancellation coefficients. All of the echo cancellation coefficients are set to zero at step 500. The initial characterisation of the transmission medium can be performed either on-line by a general purpose processor within the voice processing system or

off-line using a general purpose computer. Step 505 outputs a white noise signal over the transmission medium. Step 510 calculates an estimate, $y'(k)$, of the received echo signal using the current echo cancellation coefficients, $h(k)$. An incoming signal, $y(k)$, is received at step 515. Step 520 calculates an error signal, $e(k)$, which is the difference between the received echo signal, $y(k)$, and the estimate, $y'(k)$, thereof. The error signal, $e(k)$, is used, at step 525, to iteratively modify all the coefficients of the set of echo cancellation coefficients as follows:

$$h_i(k+1) = h_i(k) - \alpha \cdot e(k) \cdot x(k-D-i+1), \text{ for } i = 1 \text{ to } N,$$

where $h_i()$ represents the i th coefficient,

N represents the filter length, and

k represent the k th estimate of the set of echo cancellation coefficients.

The echo cancellation coefficients may then stored or output for further processing. However, it is preferable that the echo cancellation coefficients are further refined. Hence, a output signal, $x(k)$, having a suitable length is utilised and the refinement of the echo cancellation coefficients is continued for the duration of the training sequence. Step 530 determines whether or not there are more samples of the output signal, $x(k)$, to be output. If so, processing continues at step 505. If not, initialisation of the echo cancellation coefficient is complete. The echo cancellation coefficients are then stored for use during subsequent echo cancellation.

As there is invariably a delay between transmission of an output signal and the receipt of an echo signal derived therefrom, a further embodiment does not commence processing of the signal present at the input to the voice processing system until a predeterminable period of time has elapsed. Allowing said predeterminable period of time to elapse ensures that the echo signal comprises echoes and not merely noise or other signals intrinsically present on the transmission medium. The magnitude of the delay before processing an echo signal is dependent upon the proximity of the source of the echoes which are to be cancelled. For example, if the predominant source of echoes is four milliseconds from the voice processing system, a delay of eight milliseconds should be utilised.

Figure 6 shows a schematic representation of voice response system 600 according to an embodiment. The output signal, $x(k)$, is output, via an output buffer 605, and a copy thereof is concurrently in a buffer or delay line 610. The taps of the delay line 610 are spaced apart by 125 microseconds. The length, N , of the filter is 32. The output signal, $x(k)$, is transmitted over a link 615 to a local channel bank 620. The local channel bank 620 contains hybrids which are a source of echoes. Any echoes generated within the channel bank 620 are propagated back to the voice response system 600 via a link therebetween 625. The echo signal, $y(k)$, which may comprise both echoes and a signal produced by a caller, is received and stored in an input buffer 630. A filter 635 calculates, substantially concurrently with said output, an estimate of the echo signal, $y'(k)$, using the echo coefficients, $h(k)$, the copy of the outgoing signal stored in the buffer or delay line 610 and the formula

$$y'(k) = h(1)x(k-D) + h(2)x(k-D-1) + \dots + h(N)x(k-D+N-1).$$

The error signal, $e(k)$, representing the difference between the echo signal, $y(k)$, and the estimate of the echo signal, $y'(k)$, is calculated via suitable arithmetic means 640, such as an arithmetic unit or a DSP. The error signal, $e(k)$, is used to adaptively modify the echo cancellation coefficient according to the following formula:

$$h_i(k+1) = h_i(k) - \alpha \cdot e(k) \cdot x(k-D-i+1), \text{ for } i = 1 \text{ to } N,$$

where $h_i()$ represents the i th coefficient,
 N represents the filter length, and
 k represent the k th estimate of the set of echo cancellation coefficients.

It can be seen from the above equation that the first set of echo cancellation coefficients generated will be $h_i(1)$ at discrete time $k = 1$, $i = 1$ to N , the second, and more refined, set of echo cancellation coefficients will be $h_i(k)$ at time $k = 2$ etc.

The modification means 645 may comprise an arithmetic unit or the DSP or may be the same arithmetic means 640 as that which performed the subtraction above. The error signal, $e(k)$, is then output for further processing by the voice processing system.

Although the above embodiment is primarily concerned with reducing the impact of locally generated echoes, the present invention can be utilised on a per call basis. However, the primary advantage realised by the invention resides in the simple initialisation step of making precalculated echo cancellation coefficients available for use in echo cancellation thereby reducing the processing load of the DSPs within a voice processing system.

However, if local characterisation is required on a per call basis, the initial determination of echo cancellation coefficients can be used to calculate the echo cancellation coefficients on a per call basis.

In a further embodiment, the error signal, $e(k)$, can be used to adaptively modify the echo cancellation coefficients during the call. The echo cancellation coefficients are modified according to the following:

$$h_i(k+1) = h_i(k) - \alpha \cdot e(k) \cdot x(k-D-i+1), \text{ for } i = 1 \text{ to } N,$$

where $h_i()$ represents the i th coefficient,
 N represents the filter length, and
 k represent the k th estimate of the set of echo cancellation coefficients.

The previously calculated echo cancellation coefficients provide a good starting point from which the characterisation, on a per basis, of the transmission medium can be performed. Characterising the transmission medium using pre-existing echo cancellation coefficients facilitates more rapid convergence of those coefficients to stable echo cancellation coefficients.

CLAIMS

1. A method for performing echo cancellation, within a voice processing system during transmission of an audio signal over a communication network, said communication network comprising a local source of echoes, said method comprising the steps of

initially determining a set of echo cancellation coefficients suitable for cancelling locally generated echoes, and

performing echo cancellation using said set of echo cancellation coefficients for all subsequent transmissions over said communication network.

2. A method as claimed in claim 1, wherein said step of initially determining a set of echo cancellation coefficients comprises the steps of

(a) setting all echo cancellation coefficients to a value of zero,

(b) transmitting an output signal $(x(k))$ to said transmission medium (515),

(c) storing a copy of said output signal $(x(k))$ in a buffer,

(d) receiving from said transmission medium (525) an echo signal $(y(k))$ representative of a locally distorted version of said output signal $(x(k))$,

(e) generating an estimate $(y'(k))$ of said echo signal using said copy of said output signal, and

(f) modifying said echo cancellation coefficients according to the difference between said echo signal $(x(k))$ and said estimate $(y'(k))$ thereof.

3. A method as claimed in either of claims 1 or 2, further comprising the step of

repeating said steps (b) to (f) until said modified echo cancellation coefficients are stable.

4. A method as claimed in claim 3, wherein said step of modifying comprises the steps of

selecting a suitable convergence factor, and

subtracting from an echo cancellation coefficient $h(k)$ the product of a convergence factor (α), a current estimate of said echo signal ($e(k)$) and a selectable sample ($x(k-D-i+1)$) of said copy of said output signal.

5. A method as claimed in claim 4, wherein the step of selecting a suitable convergence factor (α) comprises

establishing an initial convergence factor (α_i) which would produce very slow convergence, and

setting said suitable convergence factor (α) to approximately one quarter of the value of the initial convergence factor.

6. A method as claimed in any preceding claim, further comprising the step of

delaying sampling of said echo signal by a predeterminable period of time.

7. A method as claimed in claim 6, wherein said predeterminable period of time is derived from the time taken to condition the outgoing signal ($y(k)$) for transmission over the network.

8. A method as claimed in any preceding claim, wherein said step of performing echo cancellation using said set of echo cancellation coefficients comprises the steps of

(a) transmitting an output signal ($x(k)$) to said transmission medium (515),

(b) storing a copy of said output signal ($x(k)$) in a buffer,

(c) receiving from said transmission medium (525) an incoming signal ($y(k)$) comprising at least an echo signal representative of a locally distorted version of said output signal ($x(k)$),

- (d) generating an estimate ($y'(k)$) of said echo signal using said copy of said output signal ($x(k)$),
- (e) modifying said incoming signal ($y(k)$) by subtracting therefrom said estimate of said echo signal ($y'(k)$), and
- (f) outputting said modified signal for further processing.

9. A method as claimed in claim 8, where said further processing comprises determining whether or not said modified signal represents a DTMF tone.

10. A method as claimed in either of claims 8 or 9, further comprising the step of modifying said echo cancellation coefficients ($h(k)$) according to the difference between said incoming signal ($x(k)$) and said estimate of said echo signal ($y'(k)$).

11. A system for performing echo cancellation, within a voice processing system (160) during transmission of an audio signal over a communication network (130) between said audio signal and a telephone (325), said communication network comprising a local source of echoes (150), said system comprising

means (310) for initially determining a set of echo cancellation coefficients ($h(k)$) suitable for cancelling locally generated echoes, and

means (170) for performing echo cancellation using said set of echo cancellation coefficients for all subsequent transmissions over said communication network.

12. A system as claimed in claim 11, wherein said means for initially determining a set of echo cancellation coefficients comprises

means for setting all echo cancellation coefficients to a value of zero,

means for transmitting an output signal ($x(k)$) to said transmission medium (515),

means for storing a copy of said output signal ($x(k)$) in a buffer,

means for receiving from said transmission medium (525) an echo signal ($y(k)$) representative of a locally distorted version of said output signal ($x(k)$),

5 means for generating an estimate ($y'(k)$) of said echo signal using said copy of said output signal, and

means for modifying said echo cancellation coefficients according to the difference between said echo signal ($x(k)$) and said estimate ($y'(k)$) thereof.

10

13. A system as claimed in claim 12, further comprising

means for repeatedly executing said means for transmitting, storing, receiving, generating and modifying until said modified echo cancellation coefficients are stable.

15

14. A system as claimed in any of claims 12 to 13, wherein said means for modifying

20

means for selecting a suitable convergence factor, and

means for subtracting from an echo cancellation coefficient $h(i)$ the product of a convergence factor (α), a current of said estimate of said echo signal ($e(k)$) and a selectable sample ($x(k-D-i+1)$) of said copy of said output signal.

25

15. A system as claimed any of claims 12 to 14, wherein the means for selecting a suitable convergence factor (α) comprises

30

means for establishing an initial convergence factor (α_0) which would produce very slow convergence, and

means for setting said suitable convergence factor (α) to one quarter of the value of the initial convergence factor.

35

16. A system as claimed in any of claims 12 to 15, wherein the means for modifying further comprises

means for delaying sampling of said echo signal by a predeterminable period of time.

40

17. A system as claimed in claim 16, wherein said predeterminable period of time is derived from the time taken to condition the outgoing signal $(y(k))$ for transmission over the network.

5 18. A system as claimed in any of claims 12 to 17, wherein said means for performing echo cancellation using said set of echo cancellation coefficients comprises

10 means for transmitting an output signal $(x(k))$ to said transmission medium (515),

means for storing a copy of said output signal $(x(k))$ in a buffer,

15 means for receiving from said transmission medium (525) an incoming signal $(y(k))$ comprising at least an echo signal representative of a locally distorted version of said output signal $(x(k))$,

20 means for generating an estimate $(y'(k))$ of said echo signal using said copy of said output signal $(x(k))$,

means for modifying said incoming signal $(y(k))$ by subtracting therefrom said estimate of said echo signal $(y'(k))$, and

25 means for outputting said modified signal for further processing.

19. A system as claimed in claim 18, further comprising means for modifying said echo cancellation coefficients $(h(k))$ according to the difference between said incoming signal $(y(k))$ and said estimate of said echo signal $(y'(k))$ thereof.



Application No: GB 9525791.1
Claims searched: ALL

Examiner: Mr.SAT SATKURUNATH
Date of 11 March 1996
search:

Patents Act 1977
Search Report under Section 17

Databases searched:

UK Patent Office collections, including GB, EP, WO & US patent specifications, in:

UK Cl (Ed.O): H4R: RLES, RLS

Int Cl (Ed.6): H04B

Other: Online: WPI, JAPIO, INSPEC

Documents considered to be relevant:

Category	Identity of document and relevant passage	Relevant to claims
A	GB2109209 A WESTERN - see figures 3, 4	1, 11
A	GB2109207 A WESTERN - see figures	1, 11

X Document indicating lack of novelty or inventive step
Y Document indicating lack of inventive step if combined with one or more other documents of same category.

& Member of the same patent family

A Document indicating technological background and/or state of the art.
P Document published on or after the declared priority date but before the filing date of this invention.

E Patent document published on or after, but with priority date earlier than, the filing date of this application.

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☒ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

This Page Blank (uspto)